

J. Seharaseyn



1600

## RAW SEQUENCE LISTING

DATE: 09/17/2003

PATENT APPLICATION: US/09/939,226B

TIME: 15:35:37

Input Set : A:\PTO.PG.txt

Output Set: N:\CRF4\09172003\I939226B.raw

3 <110> APPLICANT: Euroscreen S.A.  
 4 SAMSON, Michel  
 5 PARMENTIER, Marc  
 6 VASSART, Gilbert  
 7 LIBERT, Frederick  
 9 <120> TITLE OF INVENTION: Methods for Identifying Compounds which Bind the Active CCR5  
 Chemokine  
 10 Receptor  
 12 <130> FILE REFERENCE: 9409/2023C  
 14 <140> CURRENT APPLICATION NUMBER: 09/939,226B  
 15 <141> CURRENT FILING DATE: 2001-08-24  
 17 <150> PRIOR APPLICATION NUMBER: US 08/833,752  
 18 <151> PRIOR FILING DATE: 1997-04-09  
 20 <150> PRIOR APPLICATION NUMBER: US 09/626,939  
 21 <151> PRIOR FILING DATE: 2000-07-27  
 23 <150> PRIOR APPLICATION NUMBER: US 08/810,028  
 24 <151> PRIOR FILING DATE: 1997-03-04  
 26 <160> NUMBER OF SEQ ID NOS: 18  
 28 <170> SOFTWARE: PatentIn version 3.1  
 30 <210> SEQ ID NO: 1  
 31 <211> LENGTH: 792  
 32 <212> TYPE: DNA  
 33 <213> ORGANISM: Homo sapiens  
 35 <400> SEQUENCE: 1  
 36 gaattccccc aacagagcca agctctccat ctagtggaca gggaagctag cagcaaacct 60  
 38 tcccttcact acaaaacttc attgcttggc caaaaagaga gttaattcaa tgtagacatc 120  
 40 tatgtaggca attaaaaacc tattgatgta taaaacagtt tgcattcatg gagggcaact 180  
 42 aaatacatc taggacttta taaaagatca ctttttattt atgcacaggg tgggaacaaga 240  
 44 tggattatca agtggtcaagt ccaatctatg acatcaatta ttatacatcg gagccctgcc 300  
 46 aaaaaatcaa tgtgaagcaa atcgcagccc gcctcctgcc tccgctctac tcaactggtgt 360  
 48 tcatctttgg ttttgtgggc aacatgctgg tcatcctcat cctgataaac tgcaaaaggc 420  
 50 tgaagagcat gactgacatc tacctgctca acctggccat ctctgacctg tttttccttc 480  
 52 ttactgtccc ctctctgggt cactatgctg ccgcccagtg ggactttgga aatacaatgt 540  
 54 gtcaactcct gacagggctc tattttatag gcttcttctc tgggaatcttc ttcacatcc 600  
 56 tcctgacaat cgataggtac ctggctgtcg tccatgctgt gtttgcttta aaagccagga 660  
 58 cggtcacatt tgggggtgtg acaagtgtga tcacttgggt ggtggctgtg tttgcgtctc 720  
 60 tcccaggaat catctttacc agatctcaaa aagaaggtct tcattacacc tgcagctctc 780  
 62 attttcata ca 792  
 65 <210> SEQ ID NO: 2  
 66 <211> LENGTH: 1477  
 67 <212> TYPE: DNA  
 68 <213> ORGANISM: Homo sapiens  
 70 <220> FEATURE:  
 71 <221> NAME/KEY: misc\_feature

ENTERED

P. 6

## RAW SEQUENCE LISTING

DATE: 09/17/2003

PATENT APPLICATION: US/09/939,226B

TIME: 15:35:37

Input Set : A:\PTO.PG.txt

Output Set: N:\CRF4\09172003\I939226B.raw

72 &lt;222&gt; LOCATION: (1377)..(1377)

73 &lt;223&gt; OTHER INFORMATION: Any nucleotide

76 &lt;220&gt; FEATURE:

77 &lt;221&gt; NAME/KEY: misc\_feature

78 &lt;222&gt; LOCATION: (1384)..(1385)

79 &lt;223&gt; OTHER INFORMATION: Any nucleotide

82 &lt;400&gt; SEQUENCE: 2

```

83 gaattccccc aacagagcca agctctccat ctagtggaca gggaaagctag cagcaaacct    60
85 tcccttcact acaaaacttc attgcttggc caaaaagaga gttaattcaa tgtagacatc    120
87 tatgtaggca attaaaaacc tattgatgta taaaacagtt tgcattcatg gagggcaact    180
89 aaatacattc taggacttta taaaagatca ctttttattt atgcacaggg tggacaaga    240
91 tggattatca agtgtcaagt ccaatctatg acatcaatta ttatacatcg gagccctgcc    300
93 aaaaaatcaa tgtgaagcaa atcgagccc gcctcctgcc tccgctctac tcaactggtgt    360
95 tcatctttgg ttttgtgggc aacatgctgg tcatcctcat cctgataaac tgcaaaaggc    420
97 tgaagagcat gactgacatc tacctgctca acctggccat ctctgacctg tttttccttc    480
99 ttactgtccc cttctgggct cactatgctg ccgcccagtg ggactttgga aatacaatgt    540
101 gtcaactctt gacagggctc tttttatag gcttcttctc tggaatcttc ttcacatcc    600
103 tcctgacaat cgataggtac ctggctgtcg tccatgctgt gtttgcttta aaagccagga    660
105 cggtcacctt tggggtggtg acaagtgtga tcacttgggt ggtggctgtg tttgcgtctc    720
107 tcccaggaat catctttacc agatctcaaa aagaaggctc tcattacacc tgcagctctc    780
109 attttccata cagtcagtat caattctgga agaatticca gacattaaag atagtcatct    840
111 tggggctggt cctgccgctg cttgtcatgg tcatctgcta ctcggaatc ctaaaaactc    900
113 tgcttcggtg tcgaaatgag aagaagaggc acagggctgt gaggcttatc ttcacatca    960
115 tgattgttta ttttctcttc tgggctccct acaacattgt ccttctcctg aacaccttcc    1020
117 aggaattctt tggcctgaat aattgcagta gctctaacag gttggacca gctatgcagg    1080
119 tgacagagac tcttgggatg acgcactgct gcatcaaccc catcatctat gcctttgtcg    1140
121 gggagaagtt cagaaactac ctcttagtct tcttccaaaa gcacattgcc aaacgcttct    1200
123 gcaaatgctg ttctattttc cagcaagagg ctcccagcgc agcaagctca gtttacaccc    1260
125 gatccactgg ggagcaggaa atatctgtgg gcttgtgaca cggactcaag tgggctggtg    1320
W--> 127 acccagtcag agttgtgcac atggcttagt tttcatacac agcctggggt gggggtnnggt    1380
129 tggnnagagt cttttttaa aggaagttac tgttatagag ggtctaagat tcatccattt    1440
131 atttggcatc tgtttaaagt agattagatc cgaattc    1477

```

134 &lt;210&gt; SEQ ID NO: 3

135 &lt;211&gt; LENGTH: 1442

136 &lt;212&gt; TYPE: DNA

137 &lt;213&gt; ORGANISM: Homo sapiens

139 &lt;400&gt; SEQUENCE: 3

```

140 gaattccccc aacagagcca agctctccat ctagtggaca gggaaagctag cagcaaacct    60
142 tcccttcact acaaaacttc attgcttggc caaaaagaga gttaattcaa tgtagacatc    120
144 tatgtaggca attaaaaacc tattgatgta taaaacagtt tgcattcatg gagggcaact    180
146 aaatacattc taggacttta taaaagatca ctttttattt atgcacaggg tggacaaga    240
148 tggattatca agtgtcaagt ccaatctatg acatcaatta ttatacatcg gagccctgcc    300
150 aaaaaatcaa tgtgaagcaa atcgagccc gcctcctgcc tccgctctac tcaactggtgt    360
152 tcatctttgg ttttgtgggc aacatgctgg tcatcctcat cctgataaac tgcaaaaggc    420
154 tgaagagcat gactgacatc tacctgctca acctggccat ctctgacctg tttttccttc    480
156 ttactgtccc cttctgggct cactatgctg ccgcccagtg ggactttgga aatacaatgt    540
158 gtcaactctt gacagggctc tttttatag gcttcttctc tggaatcttc ttcacatcc    600
160 tcctgacaat cgataggtac ctggctgtcg tccatgctgt gtttgcttta aaagccagga    660
162 cggtcacctt tggggtggtg acaagtgtga tcacttgggt ggtggctgtg tttgcgtctc    720

```

## RAW SEQUENCE LISTING

DATE: 09/17/2003

PATENT APPLICATION: US/09/939,226B

TIME: 15:35:37

Input Set : A:\PTO.PG.txt

Output Set: N:\CRF4\09172003\I939226B.raw

```

164 tcccaggaat catctttacc agatctcaaa aagaaggctct tcattacacc tgcagctctc 780
166 attttccata cattaagat agtcatcttg gggctgggtcc tgccgctgct tgtcatggct 840
168 atctgctact cgggaatcct aaaaactctg cttcgggtgct gaaatgagaa gaagaggcac 900
170 agggctgtga ggcttatctt caccatcatg attgtttatt ttctcttctg ggctccctac 960
172 aacattgtcc ttctcctgaa caccttccag gaattctttg gcctgaataa ttgcagtagc 1020
174 tctaacaggt tggaccaagc tatgcagggtg acagagactc ttgggatgac gcaactgctgc 1080
176 atcaacccca tcatctatgc ctttgtcggg gagaagttca gaaactacct cttagtcttc 1140
178 ttccaaaagc acattgcca acgcttctgc aaatgctgtt ctattttcca gcaagaggct 1200
180 cccgagcgag caagctcagt ttacaccga tccactgggg agcaggaaat atctgtgggc 1260
182 ttgtgacacg gactcaagtg ggctgggtgac ccagtcagag ttgtgcacat ggcttagttt 1320
184 tcatacacag cctgggctgg ggggtggttg gaggtctttt ttaaaaggaa gttactgtta 1380
186 tagagggtct aagattcatc catttatctg gcatctgttt aaagtagatt agatccgaat 1440
188 tc 1442

```

191 &lt;210&gt; SEQ ID NO: 4

192 &lt;211&gt; LENGTH: 184

193 &lt;212&gt; TYPE: PRT

194 &lt;213&gt; ORGANISM: Homo sapiens

196 &lt;400&gt; SEQUENCE: 4

```

198 Met Asp Tyr Gln Val Ser Ser Pro Ile Tyr Asp Ile Asn Tyr Tyr Thr
199 1 5 10 15
202 Ser Glu Pro Cys Gln Lys Ile Asn Val Lys Gln Ile Ala Ala Arg Leu
203 20 25 30
206 Leu Pro Pro Leu Tyr Ser Leu Val Phe Ile Phe Gly Phe Val Gly Asn
207 35 40 45
210 Met Leu Val Ile Leu Ile Leu Ile Asn Cys Lys Arg Leu Lys Ser Met
211 50 55 60
214 Thr Asp Ile Tyr Leu Leu Asn Leu Ala Ile Ser Asp Leu Phe Phe Leu
215 65 70 75 80
218 Leu Thr Val Pro Phe Trp Ala His Tyr Ala Ala Ala Gln Trp Asp Phe
219 85 90 95
222 Gly Asn Thr Met Cys Gln Leu Leu Thr Gly Leu Tyr Phe Ile Gly Phe
223 100 105 110
226 Phe Ser Gly Ile Phe Phe Ile Ile Leu Leu Thr Ile Asp Arg Tyr Leu
227 115 120 125
230 Ala Val Val His Ala Val Phe Ala Leu Lys Ala Arg Thr Val Thr Phe
231 130 135 140
234 Gly Val Val Thr Ser Val Ile Thr Trp Val Val Ala Val Phe Ala Ser
235 145 150 155 160
238 Leu Pro Gly Ile Ile Phe Thr Arg Ser Gln Lys Glu Gly Leu His Tyr
239 165 170 175
242 Thr Cys Ser Ser His Phe Pro Tyr
243 180

```

246 &lt;210&gt; SEQ ID NO: 5

247 &lt;211&gt; LENGTH: 352

248 &lt;212&gt; TYPE: PRT

249 &lt;213&gt; ORGANISM: Homo sapiens

251 &lt;400&gt; SEQUENCE: 5

```

253 Met Asp Tyr Gln Val Ser Ser Pro Ile Tyr Asp Ile Asn Tyr Tyr Thr
254 1 5 10 15

```

## RAW SEQUENCE LISTING

DATE: 09/17/2003

PATENT APPLICATION: US/09/939,226B

TIME: 15:35:37

Input Set : A:\PTO.PG.txt

Output Set: N:\CRF4\09172003\I939226B.raw

```

257 Ser Glu Pro Cys Gln Lys Ile Asn Val Lys Gln Ile Ala Ala Arg Leu
258          20          25          30
261 Leu Pro Pro Leu Tyr Ser Leu Val Phe Ile Phe Gly Phe Val Gly Asn
262          35          40          45
265 Met Leu Val Ile Leu Ile Leu Ile Asn Cys Lys Arg Leu Lys Ser Met
266          50          55          60
269 Thr Asp Ile Tyr Leu Leu Asn Leu Ala Ile Ser Asp Leu Phe Phe Leu
270 65          70          75          80
273 Leu Thr Val Pro Phe Trp Ala His Tyr Ala Ala Ala Gln Trp Asp Phe
274          85          90          95
277 Gly Asn Thr Met Cys Gln Leu Leu Thr Gly Leu Tyr Phe Ile Gly Phe
278          100         105         110
281 Phe Ser Gly Ile Phe Phe Ile Ile Leu Leu Thr Ile Asp Arg Tyr Leu
282          115         120         125
285 Ala Val Val His Ala Val Phe Ala Leu Lys Ala Arg Thr Val Thr Phe
286          130         135         140
289 Gly Val Val Thr Ser Val Ile Thr Trp Val Val Ala Val Phe Ala Ser
290 145         150         155         160
293 Leu Pro Gly Ile Ile Phe Thr Arg Ser Gln Lys Glu Gly Leu His Tyr
294          165         170         175
297 Thr Cys Ser Ser His Phe Pro Tyr Ser Gln Tyr Gln Phe Trp Lys Asn
298          180         185         190
301 Phe Gln Thr Leu Lys Ile Val Ile Leu Gly Leu Val Leu Pro Leu Leu
302          195         200         205
305 Val Met Val Ile Cys Tyr Ser Gly Ile Leu Lys Thr Leu Leu Arg Cys
306          210         215         220
309 Arg Asn Glu Lys Lys Arg His Arg Ala Val Arg Leu Ile Phe Thr Ile
310 225         230         235         240
313 Met Ile Val Tyr Phe Leu Phe Trp Ala Pro Tyr Asn Ile Val Leu Leu
314          245         250         255
317 Leu Asn Thr Phe Gln Glu Phe Phe Gly Leu Asn Asn Cys Ser Ser Ser
318          260         265         270
321 Asn Arg Leu Asp Gln Ala Met Gln Val Thr Glu Thr Leu Gly Met Thr
322          275         280         285
325 His Cys Cys Ile Asn Pro Ile Ile Tyr Ala Phe Val Gly Glu Lys Phe
326          290         295         300
329 Arg Asn Tyr Leu Leu Val Phe Phe Gln Lys His Ile Ala Lys Arg Phe
330 305         310         315         320
333 Cys Lys Cys Cys Ser Ile Phe Gln Gln Glu Ala Pro Glu Arg Ala Ser
334          325         330         335
337 Ser Val Tyr Thr Arg Ser Thr Gly Glu Gln Glu Ile Ser Val Gly Leu
338          340         345         350
341 <210> SEQ ID NO: 6
342 <211> LENGTH: 215
343 <212> TYPE: PRT
344 <213> ORGANISM: Homo sapiens
346 <400> SEQUENCE: 6
348 Met Asp Tyr Gln Val Ser Ser Pro Ile Tyr Asp Ile Asn Tyr Tyr Thr
349 1          5          10          15

```

## RAW SEQUENCE LISTING

DATE: 09/17/2003

PATENT APPLICATION: US/09/939,226B

TIME: 15:35:37

Input Set : A:\PTO.PG.txt

Output Set: N:\CRF4\09172003\I939226B.raw

```

352 Ser Glu Pro Cys Gln Lys Ile Asn Val Lys Gln Ile Ala Ala Arg Leu
353          20          25          30
356 Leu Pro Pro Leu Tyr Ser Leu Val Phe Ile Phe Gly Phe Val Gly Asn
357          35          40          45
360 Met Leu Val Ile Leu Ile Leu Ile Asn Cys Lys Arg Leu Lys Ser Met
361          50          55          60
364 Thr Asp Ile Tyr Leu Leu Asn Leu Ala Ile Ser Asp Leu Phe Phe Leu
365 65          70          75          80
368 Leu Thr Val Pro Phe Trp Ala His Tyr Ala Ala Ala Gln Trp Asp Phe
369          85          90          95
372 Gly Asn Thr Met Cys Gln Leu Leu Thr Gly Leu Tyr Phe Ile Gly Phe
373          100         105         110
376 Phe Ser Gly Ile Phe Phe Ile Ile Leu Leu Thr Ile Asp Arg Tyr Leu
377          115         120         125
380 Ala Val Val His Ala Val Phe Ala Leu Lys Ala Arg Thr Val Thr Phe
381          130         135         140
384 Gly Val Val Thr Ser Val Ile Thr Trp Val Val Ala Val Phe Ala Ser
385 145          150          155          160
388 Leu Pro Gly Ile Ile Phe Thr Arg Ser Gln Lys Glu Gly Leu His Tyr
389          165          170          175
392 Thr Cys Ser Ser His Phe Pro Tyr Ile Lys Asp Ser His Leu Gly Ala
393          180          185          190
396 Gly Pro Ala Ala Ala Cys His Gly His Leu Leu Leu Gly Asn Pro Lys
397          195          200          205
400 Asn Ser Ala Ser Val Ser Lys
401          210          215
404 <210> SEQ ID NO: 7
405 <211> LENGTH: 360
406 <212> TYPE: PRT
407 <213> ORGANISM: Homo sapiens
409 <220> FEATURE:
410 <221> NAME/KEY: MISC FEATURE
411 <222> LOCATION: (325)..(327)
412 <223> OTHER INFORMATION: Xaa = any amino acid
415 <400> SEQUENCE: 7
417 Met Leu Ser Thr Ser Arg Ser Arg Phe Ile Arg Asn Thr Asn Glu Ser
418 1          5          10          15
421 Gly Glu Glu Val Thr Thr Phe Phe Asp Tyr Asp Tyr Gly Ala Pro Cys
422          20          25          30
425 His Lys Phe Asp Val Lys Gln Ile Gly Ala Gln Leu Leu Pro Pro Leu
426          35          40          45
429 Tyr Ser Leu Val Phe Ile Phe Gly Phe Val Gly Asn Met Leu Val Val
430          50          55          60
433 Leu Ile Leu Ile Asn Cys Lys Lys Leu Lys Cys Leu Thr Asp Ile Tyr
434 65          70          75          80
437 Leu Leu Asn Leu Ala Ile Ser Asp Leu Leu Phe Ile Ile Thr Leu Pro
438          85          90          95
441 Leu Trp Ala His Ser Ala Ala Asn Glu Trp Val Phe Gly Asn Ala Met
442          100         105         110

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/09/939,226B

DATE: 09/17/2003  
TIME: 15:35:38

Input Set : A:\PTO.PG.txt  
Output Set: N:\CRF4\09172003\I939226B.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:2; N Pos. 1377,1384,1385

Seq#:7; Xaa Pos. 325,326,327

Seq#:8; Xaa Pos. 231,232,233,333,334,335

Seq#:10; Xaa Pos. 145,146,147,321,322,323

## VERIFICATION SUMMARY

DATE: 09/17/2003

PATENT APPLICATION: US/09/939,226B

TIME: 15:35:38

Input Set : A:\PTO.PG.txt

Output Set: N:\CRF4\09172003\I939226B.raw

L:127 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:1320  
M:341 Repeated in SeqNo=2  
L:497 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:320  
L:584 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:224  
M:341 Repeated in SeqNo=8  
L:774 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10 after pos.:144  
M:341 Repeated in SeqNo=10  
L:888 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:14  
L:904 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:15  
L:919 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:16  
L:934 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:17